A Syndicated Adaptive News-feed: The Energy Hub

Dinesh Nagumothu and Peter W. Eklund
dnagumot,peter.eklund@deakin.edu.au
School of Information Technology, Deakin University
Geelong, Australia

ABSTRACT

Adaptive User Interfaces (AUI) change their presentation and content depending on the individual needs of their client. In our work, we aim to create a AUI that collects articles on "energy-related" issues from various publishers using Google's Custom Search and NEWS API. We call this AUI news-feed service "The Energy Hub". We have compared the articles that result from a search from generic search engines with that of "Energy Hub". The latter shows more promising results with better precision. In future work, we aim to provide all thematically related articles to the reader, customised by user preferences and by their interactions with the AUI. The intention is to improve the reader's experience using text mining and search-term clustering, including topic modelling. An important aspect of our work is the judgement of the success of the Energy-hub AUI, and we identify a methodology for its evaluation.

CCS CONCEPTS

• **Information systems → Web search engines**; *Content ranking*; *Personalization*; *Information retrieval query processing*.

KEYWORDS

Adaptable user interface, news recommender system, topic modelling, text mining.

ACM Reference Format:
Dinesh Nagumothu and Peter W. Eklund. 2020. A Syndicated Adaptive News-feed: The Energy Hub. In *Proceedings of the Australasian Computer Science Week Multiconference (ACSW 2020), February 4–6, 2020, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3373017.3373036

1 INTRODUCTION

Adaptive User Interface (AUI) quality and performance is determined by the relevance of the search results in a syndicated customized newsfeed. If one were to use a standard generic search API, the volume of irrelevant documents would dilute the AUI success. The use of generic search engines is therefore not optimised for the purpose of providing materials of sufficient relevance for conditioning AUIs. In reality, often search results can be counter-productive, namely where "junk articles" (articles from unconfirmed or even fake news sources) appear in the search engine results with a higher ranking than reliable articles from more credible sources.

For instance, if we search for "energy" in Google as an anonymous user, advertisements from domestic energy suppliers rank higher than articles related to serious debate on the topic "Energy". On the other hand, a few crucial articles can be overlooked in the result set since the search engine may lower a document's rank due because of differences in the search terminology of the user compared to the actual content and "aboutness" [Bruza et al. 2000; Deerwester et al. 1990] of the document.

Our research involves engineering adaptive user interfaces as a syndicated news-feed and A-B testing them. The systems we engineer can be customized to specific user-preferences and re-purposed for different news-topics. The syndicated customised news-feed site forms the basis for a platform to evaluate the effectiveness of different topic modelling algorithms and the use of other techniques used in the creation of an Adaptive User Interface (AUI). We implement this approach on a real-world application that focuses on energy discussion in Australia that we call "The Energy Hub" (energy-hub) [1]. Energy-hub is currently a MVP that fetches content through a generic search API. The search results from this app are better, when compared with a general "Google" search, when used with the same queries related to energy-related topics. Our aim is to re-use this website, and rebuild it as an AUI that changes content according to user preferences and actions.

Section 2 that follows discusses the techniques and external systems that the news-feed uses. Section 3 focuses on describing the proposed approach while Section 4 and 5 illustrate the current state and evaluation of the energy-hub comparing with other search engines. Section 6 contains future ambitions for the work followed by some conclusions.

2 BACKGROUND

In many computer systems not all features of the UI are revealed because different user groups have different functional requirements or access permissions. Thus, an adaptive user interface can be a responsive interface [Gardner 2011], but it is also more than responsive, because the navigational elements in the interface will adapt to the requirements of an individual user. In many cases, AUIs will also adapt based on the environment and the end-user platform/device. However, in our case, we are most interested in adaption of the UX based on user preferences, user search history and the documents viewed or rated as relevant [Yigitbas et al. 2019].

A personalised AUI that is effective will surely enhance the user's satisfaction and experience, but how is this benefit measured? Evaluating AUIs is difficult but can be performed with user-based testing, where users are exposed to two different AUI-systems, in

---

[1] https://github.com/dineshnagumothu/ozenergyhub

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ACSW 2020, February 4–6, 2020, Melbourne, VIC, Australia
© 2020 Association for Computing Machinery.
ACM ISBN 978-1-4503-7697-6/20/02…$15.00
https://doi.org/10.1145/3373017.3373036

random order, and asked simply which they prefer. This process can be repeated for multiple AUI variants and this type of A/B testing has become a staple for web system evaluation [Molich et al. 2004].

The UI needs to adapt according to content, so information flowing into the application defines the nature of change. External systems that can crawl the Internet provide the information flow to the AUI. Google provides an easy-to-use NEWS API [Das et al. 2007; Google 2019] to fetch live news and headlines through keyword search along with options to filter over a date range, language or publisher, and then sort results according to their publication date or popularity. Alternatively, Google custom search [Gavali 2015] gives the power of configuring traditional Google search that can adapt the document ranking and has the ability to customise the user interface for one or more users and to specific web sources. The main difference between the Google NEWS and Custom Search APIs is that the former fetches documents from all over the web, while the later allows its administrator to restrict search to particular publishers. Considering the energy-hub use-case, some useful research documents, say from scholarly journals, may not appear in Google NEWS API, but can be retrieved by directing Custom Search to their sources.

Identifying document (content) themes is a crucial step in constructing AUIs. Discovering topics and mining semantics from documents can be achieved through a natural language processing technique like topic modelling [Jelodar et al. 2019]. Latent Dirichlet Allocation (LDA), a prominent unsupervised machine learning algorithm, uses probability of word occurrence and distribution in finding topic models [Blei et al. 2003]. LDA represents documents as a mixture of topics and gives impressive results when dealing with large corpora, and it also has ability to further divide into sub-topics. There have been many variants to the LDA model and we plan to compare a few more popular of these using A/B testing.

3 PROPOSED APPROACH

To make the energy-hub AUI adaptive to reader actions and choices, the underlying process of the proposed method consists of an app engine, a topic modeller, a text classifier, a relational database and a corpus of crawled documents. The process initiates with the reader making a request by simple home page visit where the navigational elements in the UI and information presented adapts to the person who logs into the energy-hub website. The reader can search for other articles as well with the provided search affordances.

The app engine is a computing resource capable of handling user initiated requests. When a reader logs into the app, its tries to fetch topic preferences of the individual user and their respective weights from its database, this can be viewed from Fig. 1 at label 3.1. These weights sort the topics based on user interest and aid us to better arrangement the presentation of the UI. The topics of interest to the user will be presented from a pick list during the registration process and these get updated automatically, based on the way the user subsequently interacts with the app. However, users don't provide input on the weights, these can be calculated via user interaction towards each document.

Recommended documents from the corpus can be retrieved through preference-based searches at 3.2 in Fig. 1. If an initial visit, the fetched documents would purely be presented based on user's (aka reader) selected interests. Readers have an option to search for other documents using keywords of their choosing. The query passed as a parameter into the Search-function (3.3 in Fig. 1) and results get back with metadata. While searching the web, the results would reorder as per user's interests. The hybrid corpus (3.4 in Fig. 1) can be formed as an amalgamation of Google Custom Search, NEWS API and the Energy-hub's own database. The documents are saved in the form of a URL in a relational database along with metadata: title, description and author name. When a search request is initiated, the app queries its own database to retrieve documents that match the search terms. After searching its own database, the app searches the web by calling a generic search API, saves the results and returns these to the user. To arrange information in the UI, and to filter unwanted content, topics from the fetched document are extracted by a topic modeller (3.5 in Fig. 1). All retrieved documents are given as an input to the topic modeller to find topics of each document. The topics are generated as clusters, by calculating the probability of word distributions and the terms with higher probability represents the topic understandable to humans. These topics can be mapped back to an individual document to find its exact theme. A generalized Latent Dirichlet Allocation can act as a model to obtain the topics. We plan to try different topic modelling algorithms and identify the better one through A/B testing.

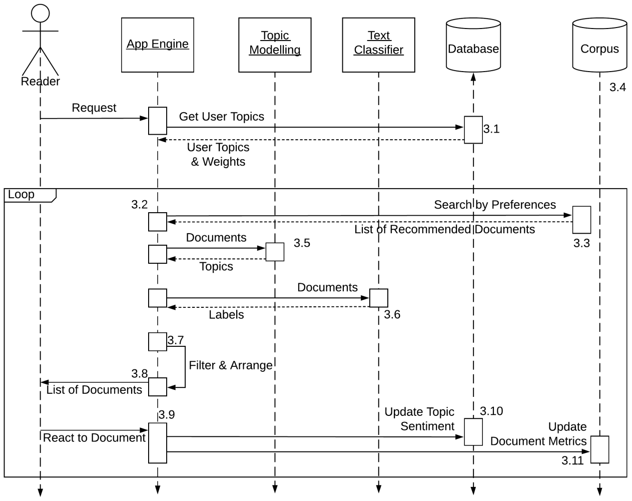


Figure 1: Sequence of events the Energy-hub approach.

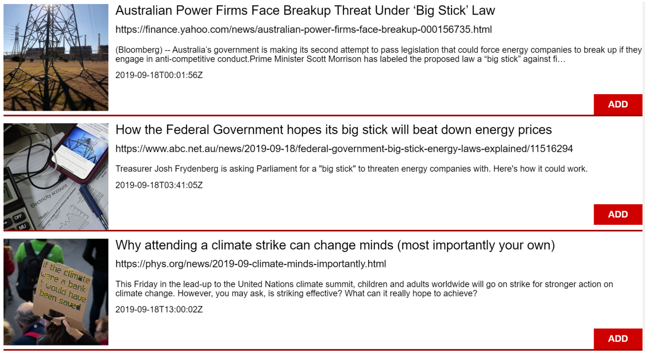
While the topic modeller reveals the relevancy of a document to the user, a text classifier (3.6 in Fig. 1) is required to classify the nature of content. A document can be a scientific journal, a news article or a blog post, but few readers in our energy-hub domain prefer searching for a single document type. One way to understand a reader's interest and arrange information is to classify each document obtained from the search result into predefined classes and attaching the predicted classes as labels to the document. A text classifier can be a machine learning model built to handle such requests by considering documents as an input and their labels as an output. After deducing topics and labels from documents, they will be compared with user interests at 3.7 in Fig. 1 to identify any irrelevant documents. User weights will be useful here to sort fetched documents as liked by the user. For a freshly registered reader, the sorting function can be skipped until weights are formed. After necessary filtering and sorting, the results can be fed back to the reader to interact with.

Before searching for articles, the app provides recommendations (3.8 in Fig. 1) based on user-defined preferences and the reader's previous interactions, along with trending and popular articles. The user can make a decision on each article to view and react based on the information provided. When the user starts to use the app, several parameters like search history, viewed articles, saved documents, their comments and ratings will be recorded to help find better content more easily in future interactions. After getting the list of recommended documents, the reader may start interacting with the system at 3.9 in Fig. 1. If the reader chooses to view, rate the document or save it to personal library, his/her sentiment towards the topics at 3.10 in Fig. 1 is recorded along with document metrics like view count at 3.11 in Fig. 1.

4 TEST CASE : THE ENERGY HUB

The "Energy Hub" (energy-hub) is interested in debate, discussion and news concerning energy production, consumption, transmission or policy in Australia. Though useful articles from Energy suppliers aren't ignored, web pages containing marketing content like Electricity and Gas retails offers are out of scope. While some documents provide meaningful information, articles with too basic information like "Facts of Coal"[2] are also tagged less useful for inclusion, the targeted readers are assumed to have prior knowledge via their declared interest.

After fetching search results, the app displays title, description and an iFrame to view the document. Registered readers may wish to save the document for later use, rate the quality, or express their opinion as comments. Popular articles from the library are displayed using the number of visits recorded per article by its readers. Forming its own corpus is an important aspect of the energy-hub, Energy related topics are published from diverse sources. News articles form an important asset as they provide daily updates concerning recent developments. Most documents in the library are collected using Google NEWS API and are constantly monitored to prevent unwanted articles. Fig. 2 displays results from Google NEWS API to a search query "energy Australia". Though the focus is on news articles, other important documents like research papers are available to the reader by search function powered by Google embedded into the energy-hub. As Google NEWS API collects daily articles, getting other types of documents, like research papers or agency reports, needs a different API. Custom search from Google can fetch other documents and act in a more directed way using pre-embedded keywords to the search engine. These terms let the search results flow in an energy-related discussion rather than a generic search outcome. Users of the energy-hub can search for articles over the web using this custom search which helps them in finding documents swiftly. The current version of the energy-hub supports news related to Australia, USA, India, Germany, Canada, UK and New Zealand as well as a composite international edition. If a country setting has changed, all search terms will update accordingly to facilitate user without having to change filters at each location.



Australian Power Firms Face Breakup Threat Under 'Big Stick' Law

https://finance.yahoo.com/news/australian-power-firms-face-breakup-000156735.html

(Bloomberg) -- Australia's government is making its second attempt to pass legislation that could force energy companies to break up if they engage in anti-competitive conduct.Prime Minister Scott Morrison has labeled the proposed law a "big stick" against fi…

2019-09-18T00:01:56Z

How the Federal Government hopes its big stick will beat down energy prices

https://www.abc.net.au/news/2019-09-18/federal-government-big-stick-energy-laws-explained/11516294

Treasurer Josh Frydenberg is asking Parliament for a "big stick" to threaten energy companies with. Here's how it could work.

2019-09-18T03:41:05Z

Why attending a climate strike can change minds (most importantly your own)

https://phys.org/news/2019-09-climate-minds-importantly.html

This Friday in the lead-up to the United Nations climate summit, children and adults worldwide will go on strike for stronger action on climate change. However, you may ask, is striking effective? What can it really hope to achieve?

2019-09-18T13:00:02Z

Figure 2: Articles retrieved by Google NEWS API for query "energy Australia".

5 EVALUATION

At this early stage, the energy-hub can be rated beneficial if it yields better results than a generic Google Search. Experiments were conducted to compare the performance of search filters applied through the energy-hub. 20 search results from both sites were rated by counting the number of relevant documents retrieved. "Smart grid energy Australia" was selected as a query to validate, as suggested by a domain expert-group. Out of 20 foremost results, the energy-hub retrieves 13 relevant documents, while traditional Google search retrieves only 8. As searching web retrieves thousands of documents, better ordering produces better results.

Precision was plotted against the document number to measure ranking performance to those 20 retrieved. It can be observed from the Fig. 3, the Google assisted energy-hub outperformed generic Google search at every position.

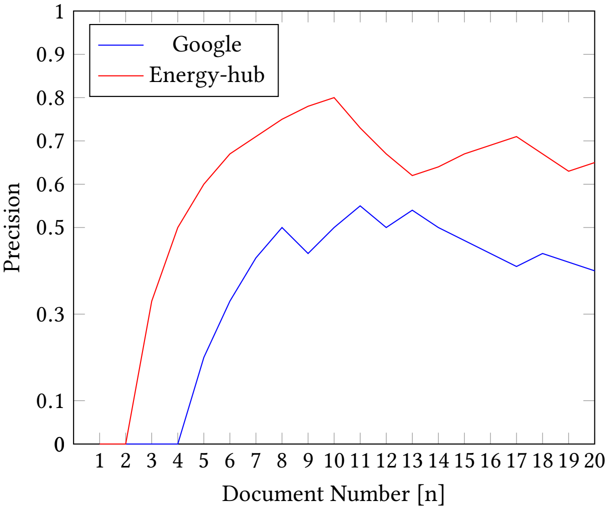


Figure 3: Comparing precision of Google vs Energy-hub for the query "Smart grid energy Australia".

As the previous query is quite specific, a different more generic query "Coal energy Australia" was used and energy-hub returned

---

[2] https://www.britannica.com/science/coal-fossil-fuel

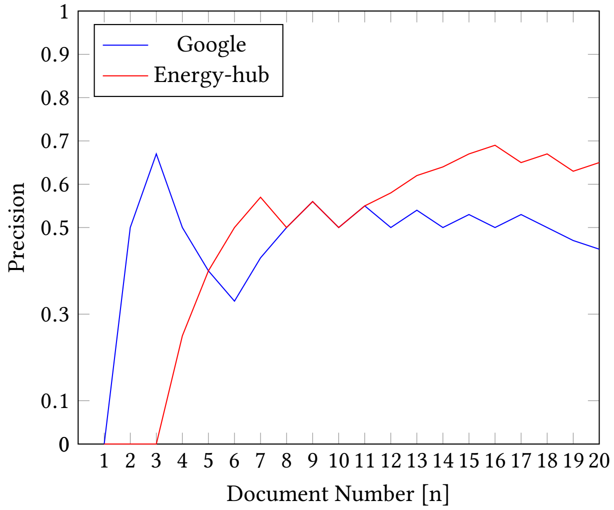


Figure 4: Comparing precision of Google vs. Energy-hub for the query "Coal energy Australia".

13 relevant documents with Google was limited to 9. Inferring from Fig. 4, unlike previous graph, precision fluctuated and Google performed well at the beginning but energy-hub successfully outperformed after the 10^{th} document. In both instances, the energy-hub performed overall better, as shown in Table 1.

Table 1: Number of high-quality documents retrieved from 20 ranked as relevant by experts over different queries.

Query terms	*Google*	*Energy-hub*
Coal energy Australia	9	13
Smart grid energy Australia	8	13

6 AMBITIONS OF ENERGY-HUB

There are a few issues with the existing system powered by Google NEWS API that need to be addressed. From our observation, it has been found that the result-set extracted from Google NEWS API contains news articles unrelated to the energy-debate. For instance, if the app searches "energy", results may include "energy drinks" or "energy in a football match" which are false friends. Currently, this process requires human intervention, as this in turn prevents the app being fully automated. Topic modelling can solve this problem by identifying the topics from a document and ensuring semantic relevance. A relevancy score can be calculated by measuring the difference between the topics that relate to the "energy" with topics of incoming documents. If the difference is higher than a threshold value, incoming documents can be tagged as irrelevant and pruned from the search. Currently, the energy-hub classifies sub-topics such as *'Grid Stability', 'Information Management', 'Coal', 'Solar'* and others. As topic modelling generates topic terms, the words can be used to tag documents, and these can be used to auto-classify documents related to a particular tag.

The intent of the energy-hub is to create a web app that can adapt to the interests of its readers. Readers register and are prompted to select the energy-related interests. This facilitates a back-end machine learning algorithm operating server-side to seamlessly provide suggested articles back to the user. The navigational elements of the energy-hub change according to reader's preferences. Initially user recommendations would be solely based on the user-selected choices but the content evolves after significant interaction with the app. Registered readers will be able to see UI updates based on their preferences, search history, article ratings and view count.

7 CONCLUSION

This paper proposes a system for creating a syndicated customized news-feed that adapts to the reader's needs. Different to other approaches, the project uses the documents previously found as a record of user "interactions" with the system and it is the contents of the corpus that influences the development of the AUI. We have included Google' Custom Search functionality, embedded required keywords framed through expert group and compared the result with "Google" general search. We experimented the framework considering two queries and energy-hub successfully outperformed both when compared to generic search. Our ultimate aim is to implement a web app through our proposed method so that the reader interface adapts to actions of the reader and content creation over the web.

Acknowledgement The Energy-hub project is supported by the Minerals Council of Australia.

REFERENCES

David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3, Jan (2003), 993–1022. http://www.jmlr.org/papers/v3/blei03a.html

Peter D Bruza, Dawei Song, Kam-Fai Wong, and Chin-Hung Cheng. 2000. Commonsense aboutness for information retrieval. *Advances in Intelligent Systems: Theory and Applications* (2000).

Abhinandan S Das, Mayur Datar, Ashutosh Garg, and Shyam Rajaram. 2007. Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th international conference on World Wide Web*. ACM, 271–280.

Scott Deerwester, Susan T Dumais, George W Furnas, Thomas K Landauer, and Richard Harshman. 1990. Indexing by latent semantic analysis. *Journal of the American society for information science* 41, 6 (1990), 391–407.

Brett S Gardner. 2011. Responsive web design: Enriching the user experience. *Sigma Journal: Inside the Digital Ecosystem* 11, 1 (2011), 13–19.

Rajashree Gavali. 2015. Discovery service for engineering and technology literature through google custom search: A case study. *DESIDOC Journal of Library & Information Technology* 35, 6 (2015).

NEWS API Google. 2019. News API - A JSON API for live news and blog articles. https://newsapi.org

Hamed Jelodar, Yongli Wang, Chi Yuan, Xia Feng, Xiahui Jiang, Yanchao Li, and Liang Zhao. 2019. Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey. *Multimedia Tools and Applications* 78, 11 (2019), 15169–15211.

Rolf Molich, Meghan R Ede, Klaus Kaasgaard, and Barbara Karyukin. 2004. Comparative usability evaluation. *Behaviour & Information Technology* 23, 1 (2004), 65–74.

Enes Yigitbas, André Hottung, Sebastian Mansfield Rojas, Anthony Anjorin, Stefan Sauer, and Gregor Engels. 2019. Context- and Data-driven Satisfaction Analysis of User Interface Adaptations Based on Instant User Feedback. *Proc. ACM Hum.-Comput. Interact.* 3, EICS, Article 19 (June 2019), 20 pages. https://doi.org/10.1145/3331161